Pervasive and Mobile Computing 23 (2015) 1–25

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Semantic web technologies in pervasive computing: A survey and research roadmap

Juan Ye [a,*], Stamatia Dasiopoulou [c], Graeme Stevenson [a], Georgios Meditskos [b], Efstratios Kontopoulos [b], Ioannis Kompatsiaris [b], Simon Dobson [a]

[a] University of St. Andrews, UK
[b] Information Technologies Institute, CERTH, Greece
[c] University Pompeu Fabra University, Barcelona, Spain

ARTICLE INFO

Article history:
Received 8 April 2014
Received in revised form 19 December 2014
Accepted 25 December 2014
Available online 9 January 2015

Keywords:
Ontologies
Pervasive computing
Streaming query
Uncertainty reasoning
Context awareness

ABSTRACT

Pervasive and sensor-driven systems are by nature open and extensible, both in terms of input and tasks they are required to perform. Data streams coming from sensors are inherently noisy, imprecise and inaccurate, with differing sampling rates and complex correlations with each other. These characteristics pose a significant challenge for traditional approaches to storing, representing, exchanging, manipulating and programming with sensor data. Semantic Web technologies provide a uniform framework for capturing these properties. Offering powerful representation facilities and reasoning techniques, these technologies are rapidly gaining attention towards facing a range of issues such as data and knowledge modelling, querying, reasoning, service discovery, privacy and provenance. This article reviews the application of the Semantic Web to pervasive and sensor-driven systems with a focus on information modelling and reasoning along with streaming data and uncertainty handling. The strengths and weaknesses of current and projected approaches are analysed and a roadmap is derived for using the Semantic Web as a platform, on which open, standard-based, pervasive, adaptive and sensor-driven systems can be deployed.

1. Introduction

Nowadays computing is pervasive, anywhere and any-time, leading to a profound impact on everyday life. To begin with, our homes have installed energy monitoring sensors and intelligent systems that automatically adjust and control heaters based on our behaviours or preferences, aiming at maximising comfort but also minimising energy consumption [1]. Furthermore, smart home based pervasive assistants can help elderly people lead independent lives by collecting symptoms-related data and adapting to different types of physical and cognitive deficits [2,3]. Even our phones have evolved into a hub of sensing, computing and communication, helping us locate featured restaurants, plan a day trip or suggest a transport mode depending on traffic conditions and weather information.

Moving beyond individuals, by collectively analysing GPS data acquired from a large number of users we can identify hot zones within a city, further contributing to tourist recommendations [4] and urban planning [5]. Outside the cities, sensors can be deployed in remote areas to monitor pollution in landfilled sites [6] or gather environmental information, in order to accurately control the concentration of fertiliser in soil [7]. Such remote sensing supports the long-term collection of fine-grained environmental data, which would be otherwise difficult or impossible to gather. This is a vital factor in

* Corresponding author.
E-mail addresses: juan.ye@st-andrews.ac.uk (J. Ye), skontopo@iti.gr (E. Kontopoulos).

http://dx.doi.org/10.1016/j.pmcj.2014.12.009
1574-1192/

fostering scientific research and increasing understanding of the environment and all these broad application scenarios are well supported by the advances in pervasive sensing and communication technologies.

On the other hand, pervasive computing faces the challenge of how to model and reason on such massive amounts of data and how to facilitate sharing and interoperability across heterogeneous systems and applications. For example, how can an intelligent traffic control system effectively use pollution data monitored in a city, in order to design a pollution-free route? Or how can a smart home energy system meaningfully use traffic information, in order to predict when a user will arrive home? Consequently, there is a pressing need for open and standards-based representations that will facilitate integrating information of heterogeneous types and modalities, as well as communicating and exchanging information between devices and components [8].

A suitable solution lies in *Semantic Web* (*SW*) technologies [9], which aim to bring to the table the ability to formally capture intended semantics and to support automated reasoning, supporting sharing, integration and management of information from heterogeneous sources. These capabilities satisfy perfectly all those common requirements in pervasive, sensor-driven and adaptive computing environments mentioned above. Via explicitly rendering meaning, the Semantic Web tries to facilitate data exchange between systems and components in an open, extensible manner, maintaining semantic clarity across applications. SW technologies have demonstrated to successfully address several pervasive computing concerns in a number of small-scaled and targeted applications, such as representing complex sensor data [10], recognising human activities [11] and modelling and querying location data across heterogeneous coordinate systems [12]. Additionally, the use of ontologies elegantly supports the cooperation of data sources within an open system [13], while ontological reasoning proves useful in manipulating structured conceptual spaces [14,8].

However, the potential of SW technologies in addressing other key pervasive computing application requirements is yet to be fully explored. Sensor data typically exhibit heterogeneous modalities and formats, real-time updating and imperfection. Such data often need to be continuously queried, aggregated to a more consistent conclusion and abstracted to different levels of generalisation for different applications types. In addition, processing and reasoning on such data are often conducted on resource-constrained devices. Thus, the key challenges include representing and reasoning on information uniformly across various sensing technologies, applications, systems and platforms; capturing temporal semantics of data and querying and applying different reasoning schemes to highly dynamic data; and, reasoning in the presence of extensive uncertainty. This survey explores these issues and seeks to answer three questions: (a) to what extent do existing SW technologies address the requirements? (b) what additional techniques might be needed? (c) how might the research community address these deficiencies?

The discussion is organised as follows. The background of SW technologies is described in Section 2. Section 3 introduces modelling information and their semantic relations at different levels of abstraction, including raw sensor data, well-structured domain information (*context*), and an atomic concept indicating a change of state (*event*). Section 4 introduces different reasoning mechanisms applied. Section 5 discusses existing strategies in modelling temporal information, facilitating querying on dynamic data and reasoning on temporal knowledge, while Section 6 discusses the approaches of handling uncertainty. Finally, Section 7 identifies challenging research issues that still require further exploration and Section 8 concludes the paper.

2. Background of Semantic Web technologies

The Semantic Web [9] is a resource-oriented extension of the current Web that aims at a common framework for sharing and reusing data across heterogeneous applications and systems. The rationale is to convert the currently unstructured and semi-structured collection of Web documents into a '*web of data*', where the underlying semantics are expressed in a formal and machine-understandable way. Within this vision, *ontologies* play a key role, providing consensual and formally-defined terms for describing resources in an unambiguous manner.

In 2004, the *Web Ontology Language* (*OWL*[1]) became a W3C recommendation, paving the way for a new generation of state of the art tools (ontology editors and reasoners) and the proliferation of ontology-based applications in several domains. Formally founded on *Description Logics* (*DL*) [15], OWL is endowed with expressive representational constructs that allow capturing complex knowledge. At the same time, OWL avails of well-defined DL reasoning services for affording automated reasoning support. These advantages furnish OWL with a variety of appealing features within the context of pervasive applications. For example, in OWL one can effectively model and reason over taxonomic knowledge. This is a desirable feature in pervasive applications, where there is the need for modelling information at different levels of granularity and abstraction that will drive the derivation of further successively detailed contexts. Similarly, OWL supports consistency checking, another useful feature when dealing with imperfect context information coming from multiple sources.

2.1. Resource Description Framework (RDF)

The *Resource Description Framework* (*RDF*[2]) provides a directed graph formalisation, with nodes representing resources and arcs representing properties. Its semantics are prescribed by two ontology languages: *RDF Schema* (*RDFS*) and OWL. RDFS

[1] http://www.w3.org/TR/owl-features/.
[2] http://www.w3.org/RDF/.